

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:27 ; Search time 299.73 Seconds  
(without alignments)  
18.379 Million cell updates/sec

Title: us-09-331-631a-7\_COPY\_34\_80  
Perfect score: 258  
Sequence: 1 YERDPRQYEQRCRCESEA.....QCEQRCEREYKEDQROOEE 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_15:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	258	100.0	525	10 Q43358	Q43358 theobroma c
2	157	60.9	625	10 Q9SPJ3	Q9SPJ3 macadamia i
3	157	60.9	666	10 Q9SPJ4	Q9SPJ4 macadamia i
4	154	59.7	666	10 Q9SPJ5	Q9SPJ5 macadamia i
5	146	56.6	593	10 Q9SEW4	Q9SEW4 juglans reg
6	95	36.8	1038	5 Q60983	Q60983 dictyostell
7	90.5	35.1	1089	12 Q40947	Q40947 kapost's sa
8	90	34.9	810	10 Q9ZM13	Q9ZM13 cucurbita m
9	89.5	34.7	1129	12 Q9OR71	Q9OR71 kapost's sa
10	88.5	34.3	393	10 Q9ZPPO	Q9ZPPO oryza sativ
11	85	32.9	551	10 Q43607	Q43607 prunus dulc
12	85	32.9	1162	12 Q98148	Q98148 kapost's sa
13	84	32.6	1108	5 Q9ND10	Q9ND10 babesia big
14	84	32.6	1339	11 Q35788	Q35788 rattus norv
15	84	32.6	2123	5 Q9UN97	Q9UN97 dictyostell
16	83.5	32.4	251	4 Q9NM29	Q9NM29 homo sapien
17	83	32.2	919	4 Q9NOST	Q9NOST homo sapien
18	82.5	32.0	554	5 Q9VPS3	Q9VPS3 drosophila
19	82	31.8	314	5 Q9VR26	Q9VR26 drosophila

20	82	31.8	608	10 Q9SUI2	Q9SUI2 arabidopsis
21	82	31.8	1027	4 Q9NS77	Q9NS77 homo sapien
22	82	31.8	1165	4 Q95819	Q95819 homo sapien
23	82	31.8	1175	4 Q75172	Q75172 homo sapien
24	82	31.8	1233	11 P97820	P97820 mus musculu
25	82	31.8	1257	4 Q95033	Q95033 homo sapien
26	82	31.8	1990	5 Q9U8Q0	Q9U8Q0 drosophila
27	82	31.8	1991	5 Q9W244	Q9W244 drosophila
28	80.5	31.2	406	2 Q87306	Q87306 borrelia bu
29	80.5	31.2	503	5 Q25777	Q25777 plasmodium
30	80	31.0	293	5 Q15988	Q15988 palinopecte
31	80	31.0	314	5 Q15987	Q15987 palinopecte
32	80	31.0	1351	5 Q96242	Q96242 plasmodium
33	79.5	30.8	411	5 P91419	P91419 caenorhabdi
34	79	30.6	199	5 Q60952	Q60952 dictyostell
35	79	30.6	304	4 Q15409	Q15409 homo sapien
36	79	30.6	517	5 Q9W4E8	Q9W4E8 drosophila
37	79	30.6	655	4 Q9NY83	Q9NY83 homo sapien
38	79	30.6	712	4 Q9NY82	Q9NY82 homo sapien
39	79	30.6	749	4 Q14244	Q14244 homo sapien
40	79	30.6	910	11 Q54899	Q54899 mus musculu
41	79	30.6	910	11 Q88704	Q88704 mus musculu
42	79	30.6	947	5 Q9XXK5	Q9XXK5 caenorhabdi
43	79	30.6	1121	4 Q94922	Q94922 homo sapien
44	78.5	30.4	948	5 Q22155	Q22155 caenorhabdi
45	78	30.2	630	5 Q9W4J3	Q9W4J3 drosophila

## ALIGNMENTS

RESULT	ID	Q43358	PRELIMINARY:	PRT:	525 AA.
AC	Q43358	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)			
DE	01-OCT-2000	(TrEMBLrel. 15, Last annotation update)			
DE	VICILIN	PRECURSOR.			
GN	CSV.				
OS	Theobroma cacao (Cacao).				
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;				
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;				
OC	Malvales; Malvaceae; Theobroma.				
OX	NCBI_TaxID=3641;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LEAVES;				
RX	MEDLINE=92288309; PubMed=1600151;				
RA	McHenry L., Fritz P.J.;				
RT	"Comparison of the structure and nucleotide sequences of vicilin genes				
RT	of cocoa and cotton raise questions about vicilin evolution."				
RL	Plant Mol. Biol. 18:1173-1176(1992).				
DR	EMBL; X62625; CAA44493.1; -.				
DR	EMBL; X62626; CAA44494.1; -.				
DR	HSSP; P02853; 2PBL.				
DR	MENDEL; 30919; Theoc.1188;30919.				
DR	INTERPRO; IPR001113; -.				
DR	PRAM; PF00546; Seedstore_7s; 1.				
DR	PRODOM; PD081059; -; 1.				
KW	Signal.				
FT	SIGNAL.				
FT	CHAIN				
FT	SEQUENCE				

Query Match 100.0%; Score 258; DB 10; Length 525;  
Best Local Similarity 100.0%; Pred. No. 5.3e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YERDPRQYEQRCRCESEATEEREDQRCEREREYKEDQROOEE 47  
DB 34 YERDPRQYEQRCRCESEATEEREDQRCEREREYKEDQROOEE 80

```
RESULT 2
09SPL3 PRELIMINARY: PRT: 625 AA.
AC 09SPL3:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR (FRAGMENT).
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia."
RL Plant J. 0:0-0(1999).
DR EMBL: AF161885; AAD54246.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
FT NON_TER 1
SQ SEQUENCE 625 AA; 73586 MW; 415808A89D370296 CRC64;
```

```
Query Match 60.9%; Score 157; DB 10; Length 625;
Best Local Similarity 53.3%; Pred. No. 2.7e-09;
Matches 24; Conservative 13; Mismatches 8; Indels 0; Gaps 0;
```

```
RESULT 3
09SPL4 PRELIMINARY: PRT: 666 AA.
AC 09SPL4:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia."
RL Plant J. 0:0-0(1999).
DR EMBL: AF161884; AAD54245.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
FT NON_TER 1
SQ SEQUENCE 666 AA; 78243 MW; 0ECA22F8710F8A7B CRC64;
```

```
Query Match 60.9%; Score 157; DB 10; Length 666;
Best Local Similarity 53.3%; Pred. No. 2.8e-09;
Matches 24; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
```

```
0Y 2 ERPRQOYEQCCRCSEATEEREEOCEORCEREYKQEOOQEE 46
DB 119 QRPQOQYEQCCRCORHETPRHMQTCQRCERREYKQKQK 163
```

```
RESULT 4
09SPL5 PRELIMINARY: PRT: 666 AA.
AC 09SPL5:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia kernels."
RL Plant J. 0:0-0(1999).
DR EMBL: AF161883; AAD54244.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
FT NON_TER 1
SQ SEQUENCE 666 AA; 78217 MW; C752B884B2DF0224 CRC64;
```

```
Query Match 59.7%; Score 154; DB 10; Length 666;
Best Local Similarity 51.1%; Pred. No. 5.9e-09;
Matches 23; Conservative 13; Mismatches 9; Indels 0; Gaps 0;
```

```
RESULT 5
09SEW4 PRELIMINARY: PRT: 593 AA.
AC 09SEW4:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS Juglans regia (English walnut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fagales; Juglandaceae; Juglans.
OX NCBI_TaxID=51240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SUNLAND; TISSUE=SOMATIC EMBRYO LINE.
RA Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;
RT "Identification and cloning of a cDNA encoding a vicilin-like protein,
RT Jug r 2, from English walnut kernel (Juglans regia): a major food
RT allergen."
RL Submitted (MAY-1998) to the EMBL/Genbank/DDJ databases.
DR EMBL: AF060555; AAF18269.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
FT NON_TER 1
SQ SEQUENCE 593 AA; 69990 MW; 9BA127E19B18C0D8 CRC64;
```

```
Query Match 56.6%; Score 146; DB 10; Length 593;
Best Local Similarity 45.8%; Pred. No. 3.7e-08;
Matches 27; Conservative 8; Mismatches 10; Indels 14; Gaps 1;
```

```
0Y 3 RDPQOYEQCCRCSEATEEREEOCEORCEREYKQEO-----RQOEE 47
DB 119 QRPQOQYEQCCRCORHETPRHMQTCQRCERREYKQKQK 163
```

Query Match	35.18;	Score 90.5;	DB 12;	Length 1089;
Best Local Similarity	41.38;	Pred. No. 0.043;		
Matches 19; Conservative	15;	Mismatches 11;	Indels 1;	Gaps 1

DR PRINTS; PRO1217; PRICEXTENSN.  
SQ SEQUENCE 1129 AA; 131346 MW; 8F63855B45F79109 CRC64;

```

Query Match          34.7%; Score 89.5; DB 12; Length 1129;
Best Local Similarity 38.8%; Pred. No. 0.057;
Matches 19; Conservative 16; Mismatches 11; Indels 3; Gaps 1;

QY 2 ERPRROVEQC--ORCESEATEEREOECORCERREYKROOQEE 47
DB 613 QQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQ 661

RESULT 10
Q92TP0 PRELIMINARY; PRT; 393 AA.
AC Q92TP0;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE HYPOTHEMETICAL 45.3 KDA PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LOMELU;
RA Chen P.W., Chen L.J.;
RA Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF049348; AAD02494.1;
KW Hypothetical protein.
SQ SEQUENCE 393 AA; 45256 MW; DBD01934BA2F9E95 CRC64;

Query Match          34.3%; Score 88.5; DB 10; Length 393;
Best Local Similarity 34.8%; Pred. No. 0.03;
Matches 16; Conservative 14; Mismatches 9; Indels 7; Gaps 2;

QY 3 RDRROVEQCQRCESEA-TEEREQCEQRCREYKROOQEE 47
DB 34 RDRKEELRMCKKOCRWACODORLORECEBOC-----LORQOEDD 73

RESULT 11
Q43607 PRELIMINARY; PRT; 551 AA.
ID Q43607;
AC Q43607;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE PRUNIN PRECURSOR.
GN PRU1.
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Rosales; Rosaceae; Prunus.
OX NCBI_TaxID=3755;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. TEXAS; TISSUE=IMMATURE SEEDLING;
RA MEDLINE-95170003; PubMed-7865791;
RA Garcia-Mas J., Messeguer R., Arus P., Puigdomenech P.;
RT "Molecular characterization of cDNAs corresponding to genes expressed
during almond (Prunus amygdalus Batsch) seed development.";
RT Plant Mol. Biol. 27:205-210(1995).
CC -1- SUBUNIT: HEXAMER; EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A
BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A
DISULFIDE BOND (BY SIMILARITY)
CC -1- SIMILARITY: TO OTHER 11S SEED STORAGE PROTEINS (GLOBULINS).
DR EMBL; X78119; CAA55009.1;
DR INTERPRO; IPR000459;
DR PFAM; PF00190; Seedstore_11s; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW Signal.

```

```

FT SIGNAL 1 20 POTENTIAL.
SQ SEQUENCE 551 AA; 63017 MW; 70D93418A22BA8F CRC64;

Query Match          32.9%; Score 85; DB 10; Length 551;
Best Local Similarity 40.5%; Pred. No. 0.093;
Matches 17; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 6 RQYEQCQRCESEATEEREOECORCERREYKROOQEE 47
DB 123 RQEQEQERQOQOQGGGGRQOQOQEQOQERQOQGRQOQOEE 164

RESULT 12
Q98148 PRELIMINARY; PRT; 1162 AA.
ID Q98148;
AC Q98148;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ORF73 HOMOLOG.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Kaposi's sarcoma-associated herpesvirus.
OX NCBI_TaxID=37296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97048116; PubMed-8892957;
RA Cesarman E., Nador R.G., Bai F., Bohenzky R.A., Russo J.J.,
RA Moore P.S., Chang Y., Knowles D.M.;
RT "Kaposi's sarcoma-associated herpesvirus contains G protein-coupled
RT receptor and cyclin D homologs which are expressed in Kaposi's sarcoma
RT and malignant lymphoma.";
RL J. Virol. 70:8218-8223(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-97094384; PubMed-8939871;
RA Moore P.S., Bashoff C., Weiss R.A., Chang Y.;
RT "Molecular mimicry of human cytokine and cytokine response pathway
RT genes by KSHV.";
RL Science 274:1739-1744(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-97121480; PubMed-8962146;
RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Madalena D.,
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
RT (HHV8) ";
RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
RN [4]
RP SEQUENCE FROM N.A.
RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Madalena D.,
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Madalena D.,
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; U52064; AAC55944.1;
DR EMBL; U75698; AAC57158.1;
DR INTERPRO; IPR002017;
SQ SEQUENCE 1162 AA; 135213 MW; 1A72CE01C1CB081C CRC64;

Query Match          32.9%; Score 85; DB 12; Length 1162;
Best Local Similarity 32.6%; Pred. No. 0.17;
Matches 15; Conservative 17; Mismatches 14; Indels 0; Gaps 0;

QY 2 ERPRROVEQCQRCESEATEEREOECORCERREYKROOQEE 47
DB 591 QREPOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQOQDEQ 636

```



